

OIPE

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/021,657

DATE: 01/15/2002

TIME: 19:02:43

RECEIVED

FEB 06 2002

STECH CENTER 1600/2900

Input Set : A:\1148.app

Output Set: N:\CRF3\01152002\J021657.raw

3 <110> APPLICANT: ALBERTSEN, MARC C.

4 FOX, TIM

5 HUFFMAN, GARY

6 TRIMNELL, MARY

8 <120> TITLE OF INVENTION: NUCLEOTIDE SEQUENCES MEDIATING MALE FERTILITY AND

9 METHOD OF USING SAME

11 <130> FILE REFERENCE: 1148

13 <140> CURRENT APPLICATION NUMBER: US/10/021,657

14 <141> CURRENT FILING DATE: 2001-12-14

16 <160> NUMBER OF SEQ ID NOS: 7

18 <170> SOFTWARE: PatentIn Ver. 2.1

20 <210> SEQ ID NO: 1

21 <211> LENGTH: 1906

22 <212> TYPE: DNA

23 <213> ORGANISM: Zea mays

25 <220> FEATURE:

26 <221> NAME/KEY: CDS

27 <222> LOCATION: (1..1638, 1642..1767)

29 <400> SEQUENCE: 1

30 gaa ttc ggc acg agg gaa gct cac ctc acg ccg gcg acg cca tcg cca 48

31 Glu Phe Gly Thr Arg Glu Ala His Leu Thr Pro Ala Thr Pro Ser Pro

32 1 5 10 15

34 ttc ttc cca cta gca ggg cct cac aag tac atc gcg ctc ctt ctg gtt 96

35 Phe Phe Pro Leu Ala Gly Pro His Lys Tyr Ile Ala Leu Leu Leu Val

36 20 25 30

38 gtc ctc tca tgg atc ctg gtc cag agg tgg agc ctg agg aag cag aaa 144

39 Val Leu Ser Trp Ile Leu Val Gln Arg Trp Ser Leu Arg Lys Gln Lys

40 35 40 45

42 ggc ccg aga tca tgg cca gtc atc ggc gca acg gtg gag cag ctg agg 192

43 Gly Pro Arg Ser Trp Pro Val Ile Gly Ala Thr Val Glu Gln Leu Arg

44 50 55 60

46 aac tac cac cgg atg cac gac tgg ctt gtc ggg tac ctg tca cgg cac 240

47 Asn Tyr His Arg Met His Asp Trp Leu Val Gly Tyr Leu Ser Arg His

48 65 70 75 80

50 agg aca gtg acc gtc gac atg ccg ttc act tcc tac acc tac atc gct 288

51 Arg Thr Val Thr Val Asp Met Pro Phe Thr Ser Tyr Thr Tyr Ile Ala

52 85 90 95

54 gac ccg gtg aat gtc gag cat gtc ctc aag act aac ttc acc aat tac 336

55 Asp Pro Val Asn Val Glu His Val Leu Lys Thr Asn Phe Thr Asn Tyr

56 100 105 110

58 ccc aag gga atc gtg tac aga tcc tac atg gac gtg ctc ctc ggt gac 384

59 Pro Lys Gly Ile Val Tyr Arg Ser Tyr Met Asp Val Leu Leu Gly Asp

60 115 120 125

62 ggc atc ttc aac gcc gac ggc gag ctg tgg agg aag cag agg aag acg 432

63 Gly Ile Phe Asn Ala Asp Gly Glu Leu Trp Arg Lys Gln Arg Lys Thr

64 130 135 140

66 gcg agt ttc gag ttc gcc tcc aag aac ctg agg gat ttc agc gcc att 480

ENTERED

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/021,657

DATE: 01/15/2002

TIME: 19:02:43

Input Set : A:\1148.app

Output Set: N:\CRF3\01152002\J021657.raw

67	Ala	Ser	Phe	Glu	Phe	Ala	Ser	Lys	Asn	Leu	Arg	Asp	Phe	Ser	Ala	Ile	
68	145					150					155					160	
70	gtg	ttc	aga	gag	tac	tcc	ctg	aag	ctg	tcg	ggt	ata	ctg	agc	cag	gca	528
71	Val	Phe	Arg	Glu	Tyr	Ser	Leu	Lys	Leu	Ser	Gly	Ile	Leu	Ser	Gln	Ala	
72						165					170					175	
74	tcc	aag	gca	ggc	aaa	gtt	gtg	gac	atg	cag	gaa	ctt	tac	atg	agg	atg	576
75	Ser	Lys	Ala	Gly	Lys	Val	Val	Asp	Met	Gln	Glu	Leu	Tyr	Met	Arg	Met	
76						180					185					190	
78	acg	ctg	gac	tcc	atc	tgc	aag	gtt	ggg	ttc	ggg	gtc	gag	atc	ggc	acg	624
79	Thr	Leu	Asp	Ser	Ile	Cys	Lys	Val	Gly	Phe	Gly	Val	Glu	Ile	Gly	Thr	
80						195					200					205	
82	ctg	tcg	cca	gat	ctc	ccc	gag	aac	agc	ttc	gcg	cag	gcg	ttc	gat	gcc	672
83	Leu	Ser	Pro	Asp	Leu	Pro	Glu	Asn	Ser	Phe	Ala	Gln	Ala	Phe	Asp	Ala	
84						210					215					220	
86	gcc	aac	atc	atc	atc	acg	ctg	cgg	ttc	atc	gac	ccg	ctg	tgg	cgc	atc	720
87	Ala	Asn	Ile	Ile	Ile	Thr	Leu	Arg	Phe	Ile	Asp	Pro	Leu	Trp	Arg	Ile	
88	225					230					235					240	
90	aag	agg	ttc	ttc	cac	gtc	ggg	tca	gag	gcc	ctc	cta	gcg	cag	agc	atc	768
91	Lys	Arg	Phe	Phe	His	Val	Gly	Ser	Glu	Ala	Leu	Leu	Ala	Gln	Ser	Ile	
92						245					250					255	
94	aag	ctc	gtg	gac	gag	ttc	acc	tac	agc	gtg	atc	cgc	cgg	agg	aag	gcc	816
95	Lys	Leu	Val	Asp	Glu	Phe	Thr	Tyr	Ser	Val	Ile	Arg	Arg	Arg	Lys	Ala	
96						260					265					270	
98	gag	atc	gtc	gag	gtc	cgg	gcc	agc	ggc	aaa	cag	gag	aag	atg	aag	cac	864
99	Glu	Ile	Val	Glu	Val	Arg	Ala	Ser	Gly	Lys	Gln	Glu	Lys	Met	Lys	His	
100						275					280					285	
102	gac	atc	ctg	tca	cgg	ttc	atc	gag	ctg	ggc	gag	gcc	ggc	gac	gac	ggc	912
103	Asp	Ile	Leu	Ser	Arg	Phe	Ile	Glu	Leu	Gly	Glu	Ala	Gly	Asp	Asp	Gly	
104						290					295					300	
106	ggc	ggc	ttc	ggg	gac	gat	aag	agc	ctc	cgg	gac	gtg	gtg	ctc	aac	ttc	960
107	Gly	Gly	Phe	Gly	Asp	Asp	Lys	Ser	Leu	Arg	Asp	Val	Val	Leu	Asn	Phe	
108	305					310					315					320	
110	gtg	atc	gcc	ggg	cgg	gac	acg	acg	gcg	acg	acg	ctg	tcg	tgg	ttc	acg	1008
111	Val	Ile	Ala	Gly	Arg	Asp	Thr	Thr	Ala	Thr	Thr	Leu	Ser	Trp	Phe	Thr	
112						325					330					335	
114	cac	atg	gcc	atg	tcc	cac	ccg	gac	gtg	gcc	gag	aag	ctg	cgc	cgc	gag	1056
115	His	Met	Ala	Met	Ser	His	Pro	Asp	Val	Ala	Glu	Lys	Leu	Arg	Arg	Glu	
116						340					345					350	
118	ctg	tgc	gcg	ttc	gag	gcg	gag	cgc	gcg	cgc	gag	gag	ggc	gtc	acg	ctc	1104
119	Leu	Cys	Ala	Phe	Glu	Ala	Glu	Arg	Ala	Arg	Glu	Glu	Gly	Val	Thr	Leu	
120						355					360					365	
122	gtg	ctc	tgc	ggc	ggc	gct	gac	gcc	gac	gac	aag	gcg	ttc	gcc	gcc	cgc	1152
123	Val	Leu	Cys	Gly	Gly	Ala	Asp	Ala	Asp	Asp	Lys	Ala	Phe	Ala	Ala	Arg	
124						370					375					380	
126	gtg	gcg	cag	ttc	gcg	ggc	ctc	ctc	acc	tac	gac	agc	ctc	ggc	aag	ctg	1200
127	Val	Ala	Gln	Phe	Ala	Gly	Leu	Leu	Thr	Tyr	Asp	Ser	Leu	Gly	Lys	Leu	
128	385					390					395					400	
130	gtc	tac	ctc	cac	gcc	tgc	gtc	acc	gag	acg	ctc	cgc	ctg	tac	ccc	gcc	1248
131	Val	Tyr	Leu	His	Ala	Cys	Val	Thr	Glu	Thr	Leu	Arg	Leu	Tyr	Pro	Ala	

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/021,657

DATE: 01/15/2002

TIME: 19:02:43

Input Set : A:\1148.app

Output Set: N:\CRF3\01152002\J021657.raw

```

132          405          410          415
134 gtc cct cag gac ccc aag ggg atc ctg gag gac gac gtg ctg ccg gac      1296
135 Val Pro Gln Asp Pro Lys Gly Ile Leu Glu Asp Asp Val Leu Pro Asp
136          420          425          430
138 ggg acg aag gtg agg gcc ggc ggg atg gtg acg tac gtg ccc tac tcg      1344
139 Gly Thr Lys Val Arg Ala Gly Gly Met Val Thr Tyr Val Pro Tyr Ser
140          435          440          445
142 atg ggg cgg atg gag tac aac tgg ggc ccc gac gcg gcg agc ttc cgg      1392
143 Met Gly Arg Met Glu Tyr Asn Trp Gly Pro Asp Ala Ala Ser Phe Arg
144          450          455          460
146 ccg gag cgg tgg atc aac gag gat ggc gcg ttc cgc aac gcg tcg ccg      1440
147 Pro Glu Arg Trp Ile Asn Glu Asp Gly Ala Phe Arg Asn Ala Ser Pro
148 465          470          475          480
150 ttc aag ttc acg gcg ttc cag gcg ggg ccg agg atc tgc ctg ggc aag      1488
151 Phe Lys Phe Thr Ala Phe Gln Ala Gly Pro Arg Ile Cys Leu Gly Lys
152          485          490          495
154 gac tcg gcg tac ctg cag atg aag atg gcg ctg gcc atc ctc ttc cgc      1536
155 Asp Ser Ala Tyr Leu Gln Met Lys Met Ala Leu Ala Ile Leu Phe Arg
156          500          505          510
158 ttc tac agc ttc cgg ctg ctg gag ggg cac ccg gtg cag tac cgc atg      1584
159 Phe Tyr Ser Phe Arg Leu Leu Glu Gly His Pro Val Gln Tyr Arg Met
160          515          520          525
162 atg acc atc ctc tcc atg gcg cac ggc ctc aag gtc cgc gtc tct agg      1632
163 Met Thr Ile Leu Ser Met Ala His Gly Leu Lys Val Arg Val Ser Arg
164          530          535          540
166 gcc gtc tga tgt cat ggc gat ttg gat atg gat atc gtc ccg ctt aat      1680
167 Ala Val Cys His Gly Asp Leu Asp Met Asp Ile Val Pro Leu Asn
168 545          550          555
170 cca cga caa ata acg ctc gtg tta caa att tgc atg cat gca tgt aag      1728
171 Pro Arg Gln Ile Thr Leu Val Leu Gln Ile Cys Met His Ala Cys Lys
172 560          565          570          575
174 gga aag cga tgg gtt tca ttg gtg gct tgg ctt aag cct taaaaactcc      1777
175 Gly Lys Arg Trp Val Ser Leu Val Ala Trp Leu Lys Pro
176          580          585
178 gtcgggtctt gcgaaccacc acatcactag tgttttgtac tctactctc agtgggaagtg      1837
180 tagtgacagc atacaagttc atcatatata ttatcctctt tcttaaaaaa aaaaaaaaaa      1897
182 aaactcgag      1906
185 <210> SEQ ID NO: 2
186 <211> LENGTH: 588
187 <212> TYPE: PRT
188 <213> ORGANISM: Zea mays
190 <400> SEQUENCE: 2
191 Glu Phe Gly Thr Arg Glu Ala His Leu Thr Pro Ala Thr Pro Ser Pro
192 1          5          10          15
194 Phe Phe Pro Leu Ala Gly Pro His Lys Tyr Ile Ala Leu Leu Leu Val
195          20          25          30
197 Val Leu Ser Trp Ile Leu Val Gln Arg Trp Ser Leu Arg Lys Gln Lys
198          35          40          45
200 Gly Pro Arg Ser Trp Pro Val Ile Gly Ala Thr Val Glu Gln Leu Arg

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/021,657

DATE: 01/15/2002

TIME: 19:02:43

Input Set : A:\1148.app

Output Set: N:\CRF3\01152002\J021657.raw

201	50	55	60
203	Asn Tyr His Arg Met His Asp Trp Leu Val Gly Tyr Leu Ser Arg His		
204	65	70	75
206	Arg Thr Val Thr Val Asp Met Pro Phe Thr Ser Tyr Thr Tyr Ile Ala		
207		85	90
209	Asp Pro Val Asn Val Glu His Val Leu Lys Thr Asn Phe Thr Asn Tyr		
210		100	105
212	Pro Lys Gly Ile Val Tyr Arg Ser Tyr Met Asp Val Leu Leu Gly Asp		
213		115	120
215	Gly Ile Phe Asn Ala Asp Gly Glu Leu Trp Arg Lys Gln Arg Lys Thr		
216		130	135
218	Ala Ser Phe Glu Phe Ala Ser Lys Asn Leu Arg Asp Phe Ser Ala Ile		
219	145	150	155
221	Val Phe Arg Glu Tyr Ser Leu Lys Leu Ser Gly Ile Leu Ser Gln Ala		
222		165	170
224	Ser Lys Ala Gly Lys Val Val Asp Met Gln Glu Leu Tyr Met Arg Met		
225		180	185
227	Thr Leu Asp Ser Ile Cys Lys Val Gly Phe Gly Val Glu Ile Gly Thr		
228		195	200
230	Leu Ser Pro Asp Leu Pro Glu Asn Ser Phe Ala Gln Ala Phe Asp Ala		
231		210	215
233	Ala Asn Ile Ile Ile Thr Leu Arg Phe Ile Asp Pro Leu Trp Arg Ile		
234	225	230	235
236	Lys Arg Phe Phe His Val Gly Ser Glu Ala Leu Leu Ala Gln Ser Ile		
237		245	250
239	Lys Leu Val Asp Glu Phe Thr Tyr Ser Val Ile Arg Arg Arg Lys Ala		
240		260	265
242	Glu Ile Val Glu Val Arg Ala Ser Gly Lys Gln Glu Lys Met Lys His		
243		275	280
245	Asp Ile Leu Ser Arg Phe Ile Glu Leu Gly Glu Ala Gly Asp Asp Gly		
246		290	295
248	Gly Gly Phe Gly Asp Asp Lys Ser Leu Arg Asp Val Val Leu Asn Phe		
249	305	310	315
251	Val Ile Ala Gly Arg Asp Thr Thr Ala Thr Thr Leu Ser Trp Phe Thr		
252		325	330
254	His Met Ala Met Ser His Pro Asp Val Ala Glu Lys Leu Arg Arg Glu		
255		340	345
257	Leu Cys Ala Phe Glu Ala Glu Arg Ala Arg Glu Glu Gly Val Thr Leu		
258		355	360
260	Val Leu Cys Gly Gly Ala Asp Ala Asp Asp Lys Ala Phe Ala Ala Arg		
261		370	375
263	Val Ala Gln Phe Ala Gly Leu Leu Thr Tyr Asp Ser Leu Gly Lys Leu		
264	385	390	395
266	Val Tyr Leu His Ala Cys Val Thr Glu Thr Leu Arg Leu Tyr Pro Ala		
267		405	410
269	Val Pro Gln Asp Pro Lys Gly Ile Leu Glu Asp Asp Val Leu Pro Asp		
270		420	425
272	Gly Thr Lys Val Arg Ala Gly Gly Met Val Thr Tyr Val Pro Tyr Ser		
273		435	440
			445

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/021,657

DATE: 01/15/2002

TIME: 19:02:43

Input Set : A:\1148.app

Output Set: N:\CRF3\01152002\J021657.raw

```

275 Met Gly Arg Met Glu Tyr Asn Trp Gly Pro Asp Ala Ala Ser Phe Arg
276   450                               455                               460
278 Pro Glu Arg Trp Ile Asn Glu Asp Gly Ala Phe Arg Asn Ala Ser Pro
279 465                               470                               475                               480
281 Phe Lys Phe Thr Ala Phe Gln Ala Gly Pro Arg Ile Cys Leu Gly Lys
282                               485                               490                               495
284 Asp Ser Ala Tyr Leu Gln Met Lys Met Ala Leu Ala Ile Leu Phe Arg
285                               500                               505                               510
287 Phe Tyr Ser Phe Arg Leu Leu Glu Gly His Pro Val Gln Tyr Arg Met
288                               515                               520                               525
290 Met Thr Ile Leu Ser Met Ala His Gly Leu Lys Val Arg Val Ser Arg
291                               530                               535                               540
293 Ala Val Cys His Gly Asp Leu Asp Met Asp Ile Val Pro Leu Asn Pro
294 545                               550                               555                               560
296 Arg Gln Ile Thr Leu Val Leu Gln Ile Cys Met His Ala Cys Lys Gly
297                               565                               570                               575
299 Lys Arg Trp Val Ser Leu Val Ala Trp Leu Lys Pro
300                               580                               585

```

303 &lt;210&gt; SEQ ID NO: 3

304 &lt;211&gt; LENGTH: 494

305 &lt;212&gt; TYPE: DNA

306 &lt;213&gt; ORGANISM: Sorghum sp.

308 &lt;220&gt; FEATURE:

309 &lt;221&gt; NAME/KEY: modified\_base

310 &lt;222&gt; LOCATION: (1)..(494)

311 &lt;223&gt; OTHER INFORMATION: "n" bases may be a, t, c, g, other or unknown

313 &lt;400&gt; SEQUENCE: 3

```

314 ggaattcggc ttatgccgtt cacttcctac acctacatcg ctgaccgggt gaatgtcgag 60
315 catgtcctca agactaactt caccaattac cccaaggggg acgtgtacag atcctacatg 120
316 gatgtgctcc tcggtgacgg catattcaac gctgacggcg agctgtggag gaagcagagg 180
317 aagacggcga gtttcgagtt cgctccaag aacctgaggg atttcagtgc caatgttttc 240
318 agagagtact ccctgaagct gtcgggcata ctgagtcagg catccaaggc aggcaaagtt 300
W--> 319 gttgacatgc aggaacttta catgaggatg acaactggact cgatctgcaa ngttgggttc 360
W--> 320 ggggtcnana tcggcacgct gtccnccgat ctcctcgaga acagcttcnc ccaagcgttc 420
W--> 321 gatgccgcta acatcatcgt cacnctgagg ttcattccacc cncgtgtggcg catccagaag 480
W--> 322 ttcttcccn gtca                               494

```

325 &lt;210&gt; SEQ ID NO: 4

326 &lt;211&gt; LENGTH: 158

327 &lt;212&gt; TYPE: PRT

328 &lt;213&gt; ORGANISM: Sorghum sp.

330 &lt;220&gt; FEATURE:

331 &lt;221&gt; NAME/KEY: MOD\_RES

332 &lt;222&gt; LOCATION: (1)..(158)

333 &lt;223&gt; OTHER INFORMATION: "Xaa" may be any, other or unknown amino acid

335 &lt;400&gt; SEQUENCE: 4

```

336 Met Pro Phe Thr Ser Tyr Thr Tyr Ile Ala Asp Pro Val Asn Val Glu
337   1           5           10           15
339 His Val Leu Lys Thr Asn Phe Thr Asn Tyr Pro Lys Gly Asp Val Tyr
340           20           25           30

```

Use of n and/or Xaa has been detected in the Sequence Listing.  
 Review the Sequence Listing to insure a corresponding  
 explanation is presented in the <220> to <223> fields of  
 each sequence using n or Xaa.

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/021,657

DATE: 01/15/2002

TIME: 19:02:44

Input Set : A:\1148.app

Output Set: N:\CRF3\01152002\J021657.raw

L:13 M:270 C: Current Application Number differs, Replaced Current Application Number  
L:14 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:319 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3  
L:320 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3  
L:321 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3  
L:322 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3  
L:357 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4  
L:360 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4